Gaussian Mixture Model Distribution

$$p(\mathrm{s}|x) = \sum_{k=1}^{K} \pi(\mathrm{a}^k|\mathrm{x}) \prod_{t=1}^{T} \phi(s_t|\mathrm{a}^k, \mathrm{x})$$

$$\mathrm{y} = [s_1, s_2, \dots, s_T]$$

$$p(\mathrm{s}|x) = \sum_{k=1}^{K} \pi(\mathrm{a}^k|\mathrm{x}) \phi(\mathrm{s}|\mathrm{a}^k, \mathrm{x})$$

Negative Log-Likelihood Loss

Data: $\{(\mathrm{x}^m, \hat{\mathrm{s}}^m)\}_{m=1}^{M}$ $\quad \hat{\mathrm{s}}^m = [s_1^m, s_2^m, \dots, s_T^m]$

$$l(\theta) = -\sum_{m=1}^{M} \sum_{k=1}^{K} \mathbf{1}(k = \hat{k}^m) \left[\log \pi(\mathrm{a}^k|\mathrm{x}^m; \theta) + \sum_{t=1}^{T} \log \mathcal{N}(s_t^m|\mathrm{a}_t^k + \mu_t^k, \Sigma_t^k; \mathrm{x}^m; \theta) \right]$$

$$\boldsymbol{\mu^k} = [\mu_1^k, \mu_2^k, \dots, \mu_T^k]$$

$$\boldsymbol{\Sigma^k} = [\Sigma_1^k, \Sigma_2^k, \dots, \Sigma_T^k]$$

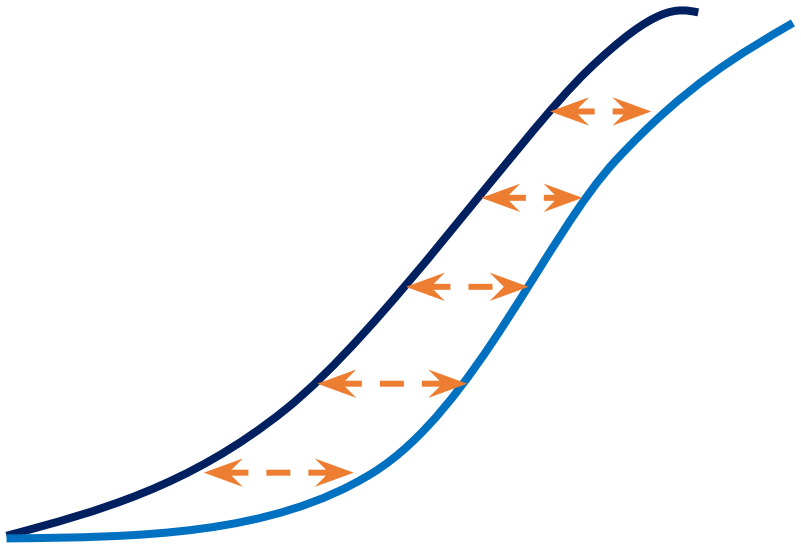
Data: $\{(\mathrm{x}^m, \mathrm{y}^m)\}_{m=1}^{M}$ $\quad \mathrm{y}^m = [y_1^m, y_2^m, \dots, y_T^m]$

$$l(\theta) = -\sum_{m=1}^{M} \sum_{k=1}^{K} \mathbf{1}(k = \hat{k}^m) \left[\log \pi(\mathrm{a}^k|\mathrm{x}^m; \theta) + \log \mathcal{N}(\mathrm{y}^m|\mathrm{a}^k + \boldsymbol{\mu^k}, \boldsymbol{\Sigma^k}; \mathrm{x}^m; \theta) \right]$$

Data: $\{(\mathrm{x}^m, \mathrm{y}^m)\}_{m=1}^{M}$

$$p(\mathrm{y}|x) = \sum_{k=1}^{K} \pi(\mathrm{a}^k|\mathrm{x})\phi(\mathrm{y}|\mathrm{a}^k, \mathrm{x})$$

$$l(\theta) = -\sum_{m=1}^{M}\sum_{k=1}^{K} \mathbf{1}\left(k = \hat{k}^m\right)\left[\log \pi(\mathrm{a}^k|\mathrm{x};\theta) + \log \mathcal{N}\left(\mathrm{y}^m|\mathrm{a}^k + \boldsymbol{\mu}^{\boldsymbol{k}}, \boldsymbol{\Sigma}^{\boldsymbol{k}}; \mathrm{x}^{\mathrm{m}}; \theta\right)\right]$$



$$\mathrm{a}^k + \mathcal{N}\left(\Delta \mathrm{s}|\mu^k, \Sigma^k\right) = y$$